# THE TOMATO Ge LOCUS: LINKAGE RELATIONS AND GEOGRAPHIC DISTRIBUTION OF ALLELES\*

# CHARLES **M.** RICK

*Department of Vegetable Crops, Uniuersity of California, Dauis, California*  Received July *27,* 1970

RASTICALLY distorted segregation of certain marker genes on chromosome  $\mathcal{L}_4$  of the tomato *(Lycopersicon esculentum Mill.)* prompted an earlier investigation (RICK 1966) of a series of alleles that effect gamete elimination. Three alleles— $Ge^n$ ,  $Ge^c$ , and  $Ge^p$ —are known at this locus, but gametes are eliminated only in  $Ge^c/Ge^p$  heterozygotes; both male and female  $Ge^c$  gametes are aborted; the interaction is approximately 95% penetrant. Segregation of linked genes is modified in such heterozygotes, the degree of disturbance depending on the tightness of the linkage. Preliminary data placed the locus of Ge in the vicinity of *ful* and *w-4,* possibly within the proximal heterochromatin of chromosome *4.* 

Studies of the  $Ge$  locus were continued in order to  $(1)$  position its locus more precisely and (2) determine which alleles are present in samples of wild and cultivated tomatoes from various geographic areas. The first objective was pursued with hope of learning more precisely what relationship, if any, the locus of  $Ge$ might have with heterochromatin and the second to ascertain what light might be shed on the relationships amongst tomato cultivars and between them and related wild species.

#### **MATERIALS AND METHODS**

*Linkage:* Since *Ge* alleles do not condition any known phenotypic effects, segregation at this locus cannot be scored directly; it is necessary instead to resort to observation *of* closely linked marker genes. Segregation of the latter can be exploited in two ways: (1) to measure linkage as a function of the degree of distortion (as in **RICK** 1966) or *(2)* to determine the allelic constitution **of** individuals by means of progeny tests. The latter method was adopted for the new studies because it gives more precise estimates, albeit at the cost of testing many progenies in an additional generation.

The precision of such tests is improved by the simultaneous segregation of several marker genes. Accordingly, a cross was made between our original stock of  $Ge^{p}$ -w-4 and *ful-Ge<sup>c</sup>-ra-e*, a new combination synthesized for this purpose. The most recent summary of linkage and cytological interrelationships is presented in Figure 1. The recessive marker genes were selected for their well-defined phenotypes expressed in early seedling development. *e* (entire) : leaves less divided and less serrate than normal. *ful* (fulgens): foliage uniformly bright yellow-green. *ra* (rava): leaves convexly recurved, elongate trichomes.  $w-4$  (wiry-4): leaf margins eroded, becoming progressively more extreme with growth. *afl* (albifolium) : cotyledons and first true leaves white, turning green.

The F, of this cross had normal phenotype, nearly 50% gamete sterility, and, after selfing, produced the large quantity of seed required for the  $\mathbf{F}_2$ . A large  $\mathbf{F}_2$  was grown and scored for the

<sup>\*</sup> **Research.supported in** part by USPHS Grant GM **06209. The assistance** of **ROMEO T. OFENA and RICHARD** W. **ZOEEI.** in **many details of this research is gratefully acknowledged.** 

## **76** *C.* **M. RICK**

## **TABLE** 1





segregation of all markers, and all  $w-4$  segregants were selected, grown to maturity, and allowed to self-pollinate for production of F, progenies. Such selection was necessary because *W-4*  is completely female sterile and its pollen production too sparse to permit large-scale testcrossing.

*Distribution of alleles:* The *Ge* allele present in a given line can be determined by crossing it with  $Ge^c$  and  $Ge^p$  testers and examining the  $F_i$ 's for signs of gamete abortion. Since various environmental stresses and certain other genetic conditions can also lead to abortion, **a** more dependable analysis is provided by F, segregations. For the purposes of the latter tests, *ful-Gee* and *GeP-w-4* testers were used. Since *w-4* homozygotes are refractory male parents and completely sterile as female parents, recently synthesized  $Ge^{p}$ -ra and  $afl$ - $Ge^{p}$  lines were used for the later matings. Cultivated and wild forms of *L. esculentum* and the closely related wild *L. pimpinellifolium* from a wide geographic range were assembled for these tests. Information on the species, area of origin, and source is given in Table 3. Several F, plants of each combination were grown and a small quantity of seed harvested from each for testing  $F<sub>2</sub>$  segregations. Normal segregation in tests against both alleles signifies the presence of the neutral *Gem* allele; normal segregation of *w-4* and a strong deficiency of *ful* detccts *Gep;* whilst normal segregation of *ful* and a great excess of *w-4* detects *Gee.* 

## **RESULTS**

Linkage: The  $F_2$  population of 2.191 individuals segregated as summarized in Table **1.** The relations typical of genes linked in a *Ge'/Gep* combination are exhibited: a huge excess **(91.5%)** of *w-4/w-4* expected from its *cis* position with  $Ge^{p}$  and deficiencies in segregations of *ful*  $(0.42\%)$ ,  $ra(0.09\%)$ , and  $e(9.8\%)$ resulting from their linkage with  $Ge^c$ . The rarity of the *ful, ra*, and *e* homozygotes obscures their mutual *cis* relationship, but sufficient *./e* segregants were recovered to reveal the *trans* relationship between *w-4* and *e.* 

As previously experienced, the precision of mapping afforded by  $F<sub>2</sub>$  tests is segregants the penetrance of *Ge* can be estimated at approximately 96%. The  $F_s$ 



**FIGURE** 1.-Genetic (above) and cytological (below) maps **of** the proximal region of chromosome *4.* Cytological locations determined by means of induced deficiencies. Centromere represented by closed oval. From KHUSH and RICK (1967).

tests outlined below and others made in the past prove beyond doubt that its action is not 100% penetrant. For the most distal locus *e,* **the** estimated distance from *Ge,* assuming 967; penetrance, is 29.7 centimorgans **(cM)** ; assuming 100% penetrance, it is 31.3 cM. For the more proximal loci, assuming 100% penetrance, the estimates for *ful, w-4*, and *ra* are 6.5, 4.4, and 3.0 cM, respectively; whilst for 96% penetrance they are 2.8, inestimable, and  $-1.1$ . Thus, although these estimates of linkage intensity correspond roughly to established values (Figure I), they are so imprecise that even the linear order is in doubt.

Much more accurate mapping is permitted by the  $F<sub>3</sub>$  data. Thanks to their unique features, these segregations yield a surprising amount of information concerning the chromosome 4 genotype of the  $F<sub>2</sub>$  parents. The high yield of a particular mutant signifies the situation of its respective gene on the same homologue with  $Ge^{p}$ , a low yield, with  $Ge^{c}$ . Linkage between markers was also readily detected, particularly when mutants segregated normally or in high yield. For these purposes progenies of 75-1 50 individuals permitted conclusive deductions, but when necessary, larger repeat progenies were grown.

In keeping with previous results, nearly all of the *w-4+* segregants were found to be  $Ge<sup>c</sup>/Ge<sup>p</sup>$ , demonstrating a substantial transmission of  $Ge<sup>c</sup>$ ; only six of the 154 tested F, individuals segregated normally. The proportion of those tested that were proven to possess *Gee* is 148/152. The tested population is set at 152 instead of 154 because the alleles could not be identified with certainty in two individuals according to the following deductions. Since there were  $187 w-4$  segregants altogether, the yield of  $Ge^c$  can be estimated at  $\frac{(148/152) \times 187}{2.191} = 0.083$ , indi-

cating 92% penetrance for the action of *Ge.* 

In the six  $\mathbf{F}_s$  progenies that segregated normally, ambiguity between a  $Ge^c/Ge^c$ and *GeP/GeP* constitution of the parent must be admitted; however, in four progenies the odds are highly in favor of the latter genotype. The situation in these four families was normal segregation for *w-4, ra,* and *e,* with *ra* and *e* in *cis* with respect to each other but in *trans* with respect to *w-4.* **A** single crossover between *ra* and *Gel'* could have yielded a homologue of *GeP-ra-e* composition, the other assumed to be a noncrossover parental *GeP-w-4.* **At** least three crossovers would have been necessary to provide *Gee* on both strands, and two very closely situated exchanges would have been required to yield a  $Ge<sup>c</sup>-w-4$  combination. Thus by argument of parsimony of assumptions, the four  $F<sub>2</sub>$  parents were much more likely to have been  $Ge^p/Ge^p$ . If this argument is accepted, it follows that  $w-4$  must lie to the right of *Ge.* For the reverse order, a *ra-Gep* recombinant would have necessarily also included  $w-4$ . Even if the unlikely  $Ge^c/Ge^c$  constitution is assumed, the required exchange between  $Ge^c$  and *ful* never included  $w-a$ , again arguing that *w-4* lies to the right of *Ge.* The gene order thus deduced *(ful-Ge* $w-4-ra-e$ ) is consonant with the crossover frequency data presented below.

The other two instances of normal segregation remain ambiguous. In both, all of the markers segregated; all ratios were Mendelian; furthermore, all linkages were in parental combinations. To explain these facts with the present model it must be assumed that crossovers close to both sides of *Ge* resulted in the exchange

#### **TABLE** *2*

$Ge^p$	Genotype				Number	Genotype					Number
		$w-4$	$+$		103	$Ge^c$					19
$Ge^{p}$	┿	$w-4$	∸	e	43	$Ge^c$				$\boldsymbol{e}$	32
Gep	ful	$w-4$			3	$Ge^c$	ful				12
Gep	ful	w <sub>4</sub>		$\epsilon$	3	$Ge^c$	ful			$\epsilon$	$^{27}$
Gev	$+$	$w-4$	ra			$Ge^c$	∸		ra		9
$Ge^{p}$	∸	$w-4$	ra	e		$Ge^c$	-	$+$	ra	$\boldsymbol{e}$	11
Ge <sup>p</sup>		$^+$	ra			$Ge^c$	ful	⊹	ra	∸	12
Ge <sup>p</sup>			ra	e	4	$Ge^c$	ful		ra	$\epsilon$	27

*Summary of gametic genotypes of 154 w-4*  $F<sub>2</sub>$  *individuals as determined by*  $F<sub>3</sub>$  *segregation* 

of *Gee* to the *w-4* homologue or *Gep* to the *ful-ra-e* homologue. No facts favor either alternative, and both are highly unlikely events. Inactivation or loss of one of the *Ge* alleles, particularly in view of the probable site of *Ge* in heterochromatin (see below), seems more plausible and is therefore arbitrarily assumed here. In any event, the final calculations are not much affected, whichever combination of events is assumed.

The 308 gametic genotypes of the 154  $w-4$  individuals thus analyzed are tallied in Table 2. The  $F_3$  tests were adequate to detect segregation of  $w$ -4 and  $e$ ; the former segregated in every progeny, and the proportions of the latter-approximately *2504* in certain families, **10%** in others-were sufficiently high to guarantee detection. In contrast, the proximally situated *ful* and *ra* segregated in such low frequencies that their presence in an  $F<sub>2</sub>$  individual could not always be ascertained by the  $F<sub>3</sub>$  tests; populations of much larger size would be required for reliable detection, especially of *ra*. Consequently, only the  $ful-Ge^p$  and  $Ge^p-ra$ crossovers could be dependably detected; whereas the reciprocals, *+-Gec* and  $Ge<sup>c</sup>$ <sub>+</sub>, could not be detected with certainty. For the purpose of estimating linkage distances of these intervals, values of the former group were therefore tallied and doubled.

The map relations thus ascertained can be represented as:



in terms of numbers of crossovers. For the region of special interest, *ful-ra,* the total length is the same **(21** crossovers) whether estimated by totalling the subintervals or by subtracting the *Ge--w-4* distance from the sum of the *ful-w-4* and *Ge-ra* distances. This part of the map can be represented thus in cM:

$$
fu/-4-6e-1.6- w-4-1.3-10-31.2-e
$$

'me agreement between these and standard values (KHUSH and **RICK** 1967) is remarkably good. The total *jul-ra* distance estimated here (6.8 cM) compares with the standard of 7. The value of **31.2** for *ra-e* is somewhat lower than the standard (as expected for single tests of large distances), but is incredibly close to the 31.3 estimated above from  $F<sub>z</sub>$  data assuming 100% penetrance and slightly higher than the 29.7 assuming 96% penetrance.

Perhaps the most important fact ascertained is that *Ge* lies between *ful* and  $w-4$ , but closer to the latter. This conclusion is in keeping with the above deductions from *Gep-ra* crossovers. When this location is considered in terms of the cytological map (Figure l), the case for implicating *Ge* with heterochromatin becomes stronger than before. The site of *ra* has been positioned in a short interval embracing the last two chromomeres of the heterochromatic zone and an equal length of the adjacent euchromatin in *4L.* Now, since *w-4* lies to the left (towards the heterochromatin) of *TU* and *Ge* lies still farther to the left, the probability of a locus in the heterochromatic zone increases. Although it is conceivable that all three loci are situated in that tiny bit of proximal euchromatin to the right of the juncture with heterochromatin, it does not seem likely because recombination fractions tend to be lower in the proximal euchromatin of the chromosomes of tomatoes and other eukaryotes (RICK 1970).

Since the crossovers presumed to have occurred between *Gep* and *ra* are crucial to the deductions regarding loci of these genes and of *w-4,* two such products were tested against  $ful-Ge^c$ . Preliminary plantings of small  $F<sub>2</sub>$  families yielded the following pooled segregations-31 +: 94 *ra*, 114 +: 1 *ful.* The segregations are disturbed precisely in the fashion of  $Ge<sup>e</sup>/Ge<sup>p</sup>$  heterozygotes, and the results verify the conclusion that *ra* and *Ger'* are linked in *cis* in the parents of both crosses.

*Geographic distribution of alleles:* Table 3 summarizes information on the source and identification of *Ge* alleles of each tested accession. In selecting temperate-zone cultivars, an attempt was made to secure the older ones of greatest historic significance. It is likely, but not certain, from collectors' data and observations of our cultures that the listed Latin American cultivars are indigenous or locally bred. The genotype is specified simply by a single symbol for each accession since duplicate tests have always yielded the same results. This situation might have been anticipated because the cultivated tomato and its close relatives are highly inbred in most environments and predominantly so in others. No evidence of heterogeneity at the *Ge* locus has yet been detected, even for the primitive lines that exhibit appreciable morphological variability (RICK 1958).

Of the 113 lines so far tested, the great majority are *L. esculentum* cultivars. Among these, nearly all the European and United States cultivars have the neutral *Ge*<sup>n</sup> allele. The exceptions are Condine (Kondine) Red and Stone with *Ge*<sup>c</sup> and the three Californian cultivars, Early Santa Clara, Pearson, and VF 6, each with  $Ge^{p}$ . The known lineal relationship between the latter three accessions undoubtedly explains their possession of this exceptional allele, although it does not suggest its origin. A distant hereditary relationship might exist between Condine Red and Stone, but my efforts to trace their pedigrees prior to 1900 have failed. Except for two instances **of** *Gec* in Central America and radical differences in the Ecuador-Peru region, all other tested wild and cultivated tomatoes registered **Gen.** 

All three alleles appear in the Peruvian material, *Gep* predominating in both wild and cultivated lines. One wild accession has  $Ge^c$ , the remaining nine,  $Ge^p$ ;

# 80 *C.* **M. RICK**

in the cultivars *Gep* is distributed from Piura in the north to Pisco and Ayacucho in the south, and in this territory several accessions **of** both *Gec* and *Ge"* were found. Except for one instance of *Gec* at a remote site in eastern Ecuador, all tested wild and cultivated accessions of that country are *Gen.* The geographic relationships of the Ecuador-Peru collections are presented in Figure 2.



__ _	
---------	--

*Source and Ge genotype of tomato accessions* 



#### **DISCUSSION**

The new experience provides extensive confirmation of the previously derived hypothesis **(RICK** 1966) concerning the nature of interaction of *Ge* alleles. The present data also verify that elimination of *Gec* gametes in *Gec/GeP* heterozygotes is not 100% effective, the new estimates for penetrance being 92 and 96% *us.* the previously derived value of 95%. The linkage relations of *Ge* are now better understood, its position, probably in the heterochromatic zone in the proximal region of chromosome 4, being well established. According to our new estimates,





Feral or wild forms of L. *esculentum,* chiefly var. *cerasiforme* 

# *L. pimpinellifolium* (wild)



Trujillo C. M. Rick LA 121 Ge<sup>C</sup><br>
Thilsa Craig, Baby Lea, Cracker Jack, Delicious, Downer's Seedling, ES 1, ES 5, Huntsman,<br>
Melville Castle, Moneymaker, Market King, Potentate, Pyports King, Radio, Scarlet Knight,<br>
Stonor



**FIGURE 2.-Distribution and genotype of tested accessions from Ecuador, Peru, southern Colombia, and northern Chile.** 

its site cannot be more than a few centimorgans away from the centromere. It is conceivable that the putative alleles might be blocks of heterochromatin, although cytological study of the parents and hybrids has not revealed any visible differences.

**A** search of the literature has not disclosed any closely parallel examples **of**  such gene action in the tomato or other higher plants. **As** to linkage relations of gametophytic factors in general, five have been situated in maize, according to the recent summary by **NEUFFER, JONES** and **ZUBER** (1968), without any apparent association with heterochromatin. It is of interest, however, that the Segregation-Distorter *(SO)* controlling element in *Drosophila melanogaster* has been located in or near the proximal heterochromatin of the right arm of chromosome *2* (SANDLER, HIRAIZUMI and SANDLER 1959), although the *SD* and *Ge*  phenomena differ considerably in the nature of their action.

As to the distribution of alleles, the most salient feature is the prevalence of *Gep* in the wild and cultivated tomatoes of Peru in contrast to the nearly exclusive existence of *Gen* elsewhere. Clearly the dominance of *Gep* in *L. pimpinellifolium*  and the *L. esculentum* cultivars must denote a close relationship between them. Either descent of the Peruvian cultivars from the former species or extensive introgression between them, implied from other evidence (RICK 1958), might explain their remarkable similarity in allelic composition.

The bearing of these findings on the origin of the cultivated tomato should be considered. Earlier speculations (RICK 1969) were made on the basis of a smaller sampling. Several factors mitigate against the usefulness of the *Ge* data for such purposes. One is the rapidity of genetic change that can take place in the artificial breeding of such highly self-pollinated, seed-propagated annuals, as witnessed by the complete replacement of tomato cultivars recently in several major production areas. Another is the problem of establishing authenticity of local varieties in Latin America. Modern communications as well as an explorative attitude of growers favor prompt exchange, even foreign introduction, of cultivars. The concordance of the Peruvian collections in respect to the unique *Gep* allele nevertheless suggests that most of them are of local origin. **A** survey of alleles at other loci controlling isozymic differences (the subject of **a** current project) might provide useful correlative evidence.

Only a single sample of var. *cerasiforme* from the Old World was included in the tests, but the *Ge"* detected therein probably prevails in this widespread weed since it is highly uniform throughout its distribution in that area. Assuming that this generalization is correct, the alien *cerasiforme* probably did not migrate from western Peru, but more likely from Mexico or Central America, where all tested accessions possess *Gen.* 

The *Ge* system offers an advantage for establishing a partial reproductive barrier. As observed previously (RICK 1966), the reduction in seed fertility of  $Ge^{c}/Ge^{p}$  heterozygotes should be sufficient to constitute a serious reproductive disadvantage. Systems of similar type with only two alleles (e.g., semisterility) suffer difficulties if they originate in sympatric populations. Thus, if subpopulation A were assumed to benefit from isolation from subpopulation B in a twoallele system, an isolating mutation would penalize matings between individuals, thereby hindering its spread in the same population. With the *Ge* system, however, mutations can take place stepwise,  $Ge^{c} \rightarrow Ge^{n} \rightarrow Ge^{p}$  or  $Ge^{c} \leftarrow Ge^{n} \rightarrow Ge^{p}$ , without affecting fertility of the heterozygotes in any single step. Thus, in the latter series, if both subpopulations A and B originally possessed *Ge", Ge"* could spread throughout A and *Gep* throughout *B* without diminishing fertility of the *Ge*  heterozygotes within subpopulations. One difficulty inherent in the three-allele system is that the mutant of the first step would not, *prima facie,* enjoy a selective advantage and would therefore require fortuitous establishment. It might be significant in respect to these speculations that both  $Ge^c$  and  $Ge^p$  exist in  $L$ . *pimpinellifolium* and all the known alleles in *L. esculentum* in contiguous, perhaps even sympatric, areas of western Peru.

# **SUMMARY**

Additional experience with the tomato gamete-eliminator locus has confirmed previous findings on the nature of allelic interaction: gametes are eliminated solely in  $Ge^c/Ge^p$  of the possible combinations of the three known alleles; male and female gametes are eliminated with equal intensity; penetrance was estimated in two tests at 92% and 96%; segregation of linked genes is distorted in proportion to the intensity of linkage with *Ge.* F, tests afforded a more precise determination of the locus of *Ge,* placing it to the right of *ful* and to the left of *w-4* on chromosome *4,* probably in the proximal heterochromatin and very close to the centromere.-The *Ge* alleles were identified in a collection of 1 **13** cultivars and wild forms of *L. esculentum* and (wild) *L. pimpinellifolium. Gen* is found throughout the tested regions of *L. esculentum* and is the exclusive or predominant allele in all areas except central and northern Peru, where the majority of accessions possess *GeP.* All three alleles were found in European and US cultivars. All tested accessions of Peruvian *L. pimpinellifolium* have *Gep* except one instance of *Ge".* The very similar allelic composition of the two species in their sympatric region in Peru suggests evolution of the cultivated forms from *L. pimpinellifohm* or extensive introgression between them. The data are inconclusive in respect to their bearing on the origin of the cultivated tomato: qualitatively they suggest a closer relationship with the cultivars of Peru, but quantitatively, with those of Central America and Mexico.

# **LITERATTJRE CITED**

JENKINS, J. A., 1948 The origin of the cultivated tomato. Econ. Botany **2:** 379-392.

- **KHUSH,** G. *S.* and C. M. RICK, 1967 Studies on the linkage map **of** chromosome *4* **of** the tomato and on the transmission of induced deficiencies. Genetica **38:** 74-94.
- NEUFFER, M. G., L. JONES and M. **S.** ZUBER, 1968 *The Muiants of Maize.* Crop Sci. Soc., Madison, Wis.
- RICK, C. M., 1958 The role of natural hybridization in the derivation of cultivated tomatoes of and on the transmission of induced deficiencies. Genetica **38:** *14*–94.<br>FFER, M. G., L. Jones and M. S. ZUBER, 1968 *The Mutants of Maize*. Crop Sci. Soc., Madi-<br>son, Wis.<br>C. C. M., 1958 The role of natural hybridization 1969 Origin of the cultivated tomato: Current status of the problem. Abstr. 12th Intern. c., C. M., 1958 The role of natural hybridization in the derivation of cultivated tomatoes of western South America. Econ. Botany 12: 346–367. ----, 1966 Abortion of male and female gametes in the tomato determined by alle plants. Proc. 2nd Stadler Symposium, Columbia, **MO.** (in press).
- SANDLER, L., Y. HIRAIZUMI and I. SANDLER, 1959 Meiotic drive in natural populations of *Drosophila mehogaster.* I: The cytogenetic basis of Segregation-Distortion. Genetics **4:**  233-250.